

AMENDMENTS TO THE CLAIMS

- 1 1. (*Currently Amended*) A computer-implemented method for providing automatic,
2 personalized information services to a user u , the method comprising:
- 3 a) transparently monitoring user interactions with data while the user is engaged in
4 normal use of a computer;
- 5 b) updating user-specific data files, wherein the user-specific data files comprise the
6 monitored user interactions with the data and a set of documents associated with
7 the user;
- 8 c) estimating parameters of a learning machine, wherein the parameters define a User
9 Model specific to the user and wherein the parameters are estimated in part from
10 the user-specific data files;
- 11 d) analyzing a document d to identify properties of the document;
- 12 e) estimating a probability $P(u|d)$ that ~~the~~ an unseen document d is of interest to the
13 user u , wherein the probability $P(u|d)$ is estimated by applying the identified
14 properties of the document to the learning machine having the parameters defined
15 by the User Model; and
- 16 f) using the estimated probability to provide automatic, personalized information
17 services to the user.
- 1 2. (*Original*) The method of claim 1 wherein the user-specific data files include
2 documents of interest to the user u and documents that are not of interest to the
3 user u , and wherein estimating the parameters comprises distinct treatment of the
4 documents of interest and the documents that are not of interest.

1 3. (*Original*) The method of claim 1 wherein analyzing the document d provides for
2 the analysis of documents having multiple distinct media types.

1 4. (*Original*) The method of claim 1 wherein transparently monitoring user
2 interactions with data comprises monitoring multiple distinct modes of user
3 interaction with network data.

1 5. (*Original*) The method of claim 4 wherein the multiple distinct modes of
2 user interaction comprise a mode selected from the group consisting of a
3 network searching mode, a network navigation mode, a network browsing
4 mode, an email reading mode, an email writing mode, a document writing
5 mode, a viewing "pushed" information mode, a finding expert advice mode,
6 and a product purchasing mode.

1 6. (*Original*) The method of claim 1 further comprising crawling network
2 documents, wherein the crawling comprises parsing crawled documents for links,
3 calculating probable user interest in the parsed links using the learning machine,
4 and preferentially following links likely to be of interest to the user.

1 7. (*Original*) The method of claim 1 wherein the identified properties of the
2 document d comprise a user u -independent property selected from the group
3 consisting of:

4 a) a probability $P(t, d)$ that the document d is of interest to users interested in a
5 topic t ;

- b) a topic classifier discrete probability distribution $P(t|d)$;
- c) a product model discrete probability distribution $P(p|d)$;
- d) product feature values extracted from the document d ;
- e) an author of the document d ;
- f) an age of the document d ;
- g) a list of documents linked to the document d ;
- h) a language of the document d ;
- i) a number of users who have accessed the document d ;
- j) a number of users who have saved the document d in a favorite document list; and
- k) a list of users previously interested in the document d .

8. (*Original*) The method of claim 1 wherein the parameters of the learning machine define a user u -dependent function selected from the group consisting of:

- a) a user topic probability distribution $P(t|u)$ representing interests of the user u in various topics t ;
- b) a user product probability distribution $P(p|u)$ representing interests of the user u in various products p ;
- c) a user product feature probability distribution $P(f|u, p)$ representing interests of the user u in various features f of each of the various products p ;
- d) a web site probability distribution $P(s|u)$ representing interests of the user u in various web sites s ;

- e) a cluster probability distribution $P(c(u)|u)$ representing similarity of the user u to users in various clusters $c(u)$;
- f) a phrase model probability distribution $P(w|u)$ representing interests of the user u in various phrases w ;
- g) an information theory based measure $I(I_w; I_u)$ representing mutual information between various phrases w and the user u ;
- h) an information theory based measure $I(I_t; I_u)$ representing mutual information between various topics t and the user u ;
- i) an information theory based measure $I(I_s; I_u)$ representing mutual information between various web sites s and the user u ;
- j) an information theory based measure $I(I_p; I_u)$ representing mutual information between various products p and the user u ; and
- k) an information theory based measure $I(I_f; I_u)$ representing mutual information between various features f of each of the various products p and the user u .

9. (Original) The method of claim 1 wherein the parameters of the learning machine define:

- a) a user product probability distribution $P(p|u)$ representing interests of the user u in various products p ; and
- b) a user product feature probability distribution $P(f|u,p)$ representing interests of the user u in various features f of each of the various products p ;

7 and wherein the method further comprises estimating a probability $P(u|d, \text{product}$
8 $\text{described}=p)$ that a document d that describes a product p is of interest to the user
9 u , wherein the probability is estimated in part from the user product probability
10 distribution and the user product feature probability distribution.

1 10. (*Original*) The method of claim 9 further comprising recommending
2 products to the user based on the probability $P(u|d, \text{product described}=p)$.

1 11. (*Original*) The method of claim 1 further comprising estimating a posterior
2 probability $P(u|d, q)$ that the document d is of interest to the user u , given a query q
3 submitted by the user.

1 12. (*Original*) The method of claim 11 wherein estimating the posterior
2 probability comprises estimating a probability $P(q|d, u)$ that the query q is
3 expressed by the user u with an information need in the document d .

1 13. (*Original*) The method of claim 1 further comprising applying the identified
2 properties of the document d to a learning machine having product parameters
3 characterizing a product p to estimate a probability $P(p|d)$ that the document d
4 refers to the product p .

- 1 14. (*Original*) The method of claim 13 further comprising updating the product
2 parameters based on the identified properties of the document d and the
3 estimated probability $P(p|d)$.
- 1 15. (*Original*) The method of claim 13 further comprising initializing the
2 product parameters based on a set of documents associated with the product
3 p .
- 1 16. (*Original*) The method of claim 1 further comprising clustering multiple users into
2 clusters of similar users, wherein the clustering comprises calculating distances
3 between User Models, and selecting similar users based on the calculated distances
4 between User Models.
- 1 17. (*Original*) The method of claim 1 further comprising calculating relative entropy
2 values between User Models of multiple users, and clustering together users based
3 on the calculated relative entropy values.
- 1 18. (*Original*) The method of claim 1 wherein the parameters defining the User Model
2 comprise calculated distances between the User Model and User Models of users
3 similar to the user.

1 19. (*Original*) The method of claim 1 further comprising selecting in a group of users
2 an expert user in an area of expertise, wherein selecting the expert user comprises
3 finding an expert User Model among User Models of the group of users, such that
4 the expert User Model indicates a strong interest of the expert user in a document
5 associated with the area of expertise.

1 20. (*Original*) The method of claim 1 further comprising parsing the document d for
2 hyperlinks, and separately estimating for each of the hyperlinks a probability that
3 the hyperlink is of interest to the user u .

1 21. (*Original*) The method of claim 1 further comprising sending to a third party web
2 server user interest information derived from the User Model, whereby the third
3 party web server may customize its interaction with the user.

1 22. (*Original*) The method of claim 1 wherein the monitored user interactions include
2 a sequence of interaction times.

1 23. (*Original*) The method of claim 1 further comprising initializing the User Model
2 using information selected from the group consisting of a set of documents
3 provided by the user, a web browser history file associated with the user, a web
4 browser bookmarks file associated with the user, ratings by the user of a set of
5 documents, and previous product purchases made by the user.

1 24. (*Original*) The method of claim 1 further comprising modifying the User Model
2 based on User Model modification requests provided by the user.

1 25. (*Original*) The method of claim 1 further comprising providing to the user a score
2 for a document identified by the user, wherein the score is derived from the
3 estimated probability.

1 26. (*Original*) The method of claim 1 further comprising providing to the user a 3D
2 map of a hyper linked document collection, wherein the 3D map indicates a user
3 interest in each document.

1 27. (*Original*) The method of claim 1 further comprising temporarily using a User
2 Model that is built from a set of predetermined parameters of a profile selected by
3 the user.

1 28. (*Original*) The method of claim 1 further comprising initializing the User Model
2 by selecting a set of predetermined parameters of a prototype user selected by the
3 user.

29. (*Original*) The method of claim 28 further comprising updating the predetermined parameters of the prototype user based on actions of users similar to the prototype user.

30. (*Original*) The method of claim 1 further comprising identifying a set of users interested in the document d .

31. (*Original*) The method of claim 30 further comprising calculating a range of interests in the document d for the identified set of users.

32. (*Currently Amended*) A program storage device accessible by a central computer, tangibly embodying a program of instructions executable by the central computer to perform method steps for providing automatic, personalized information services to a user u , the method steps comprising:

- a) transparently monitoring user interactions with data while the user is engaged in normal use of a client computer in communication with the central computer;
- b) updating user-specific data files, wherein the user-specific data files comprise the monitored user interactions with the data and a set of documents associated with the user;
- c) estimating parameters of a learning machine, wherein the parameters define a User Model specific to the user and wherein the parameters are estimated in part from the user-specific data files;

- 13 d) analyzing a document d to identify properties of the document;
- 14 e) estimating a probability $P(u|d)$ that ~~the~~ an unseen document d is of interest to the
- 15 user u , wherein the probability $P(u|d)$ is estimated by applying the identified
- 16 properties of the document to the learning machine having the parameters defined
- 17 by the User Model; and
- 18 f) using the estimated probability to provide automatic, personalized information
- 19 services to the user.

1 33. *(Original)* The program storage device of claim 32 wherein the user-specific data

2 files include documents of interest to the user u and documents that are not of

3 interest to the user u , and wherein estimating the parameters comprises distinct

4 treatment of the documents of interest and the documents that are not of interest.

1 34. *(Original)* The program storage device of claim 32 wherein analyzing the

2 document d provides for the analysis of documents having multiple distinct media

3 types.

1 35. *(Original)* The program storage device of claim 32 wherein transparently

2 monitoring user interactions with data comprises monitoring multiple distinct

3 modes of user interaction with network data.

1 36. *(Original)* The program storage device of claim 35 wherein the multiple

2 distinct modes of user interaction comprise a mode selected from the group

3 consisting of a network searching mode, a network navigation mode, a
4 network browsing mode, an email reading mode, an email writing mode, a
5 document writing mode, a viewing "pushed" information mode, a finding
6 expert advice mode, and a product purchasing mode.

1 37. (*Original*) The program storage device of claim 32 wherein the method steps
2 further comprise crawling network documents, wherein the crawling comprises
3 parsing crawled documents for links, calculating probable user interest in the
4 parsed links using the learning machine, and preferentially following links likely
5 to be of interest to the user.

1 38. (*Original*) The program storage device of claim 32 wherein the identified
2 properties of the document d comprise a user u -independent property selected
3 from the group consisting of:

- 4 a) a probability $P(t, d)$ that the document d is of interest to users interested in a
5 topic t ;
- 6 b) a topic classifier discrete probability distribution $P(t|d)$;
- 7 c) a product model discrete probability distribution $P(p|d)$;
- 8 d) product feature values extracted from the document d ;
- 9 e) an author of the document d ;
- 10 f) an age of the document d ;
- 11 g) a list of documents linked to the document d ;
- 12 h) a language of the document d ;

- 13 i) a number of users who have accessed the document d ;
- 14 j) a number of users who have saved the document d in a favorite document
- 15 list; and
- 16 k) a list of users previously interested in the document d .

- 1 39. (*Original*) The program storage device of claim 32 wherein the parameters of
- 2 the learning machine define a user u -dependent function selected from the
- 3 group consisting of:
- 4 a) a user topic probability distribution $P(t|u)$ representing interests of the user u
 - 5 in various topics t ;
 - 6 b) a user product probability distribution $P(p|u)$ representing interests of the
 - 7 user u in various products p ;
 - 8 c) a user product feature probability distribution $P(f|u, p)$ representing interests
 - 9 of the user u in various features f of each of the various products p ;
 - 10 d) a web site probability distribution $P(s|u)$ representing interests of the user u
 - 11 in various web sites s ;
 - 12 e) a cluster probability distribution $P(c(u)|u)$ representing similarity of the user
 - 13 u to users in various clusters $c(u)$;
 - 14 f) a phrase model probability distribution $P(w|u)$ representing interests of the
 - 15 user u in various phrases w ;
 - 16 g) an information theory based measure $I(I_w; I_u)$ representing mutual
 - 17 information between various phrases w and the user u ;

- h) an information theory based measure $I(I_t; I_u)$ representing mutual information between various topics t and the user u ;
- i) an information theory based measure $I(I_s; I_u)$ representing mutual information between various web sites s and the user u ;
- j) an information theory based measure $I(I_p; I_u)$ representing mutual information between various products p and the user u ; and
- k) an information theory based measure $I(I_f; I_u)$ representing mutual information between various features f of each of the various products p and the user u .

40. (*Original*) The program storage device of claim 32 wherein the parameters of the learning machine define:

- a) a user product probability distribution $P(p|u)$ representing interests of the user u in various products p ; and
 - b) a user product feature probability distribution $P(f|u, p)$ representing interests of the user u in various features f of each of the various products p ;
- and wherein the method steps further comprise estimating a probability $P(u|d, \text{product described}=p)$ that a document d that describes a product p is of interest to the user u , wherein the probability is estimated in part the user product probability distribution and the user product feature probability distribution.

1 41. (*Original*) The program storage device of claim 40 wherein the method steps
2 further comprise recommending products to the user based on the probability
3 $P(u|d, \text{product described}=p)$.

1 42. (*Original*) The program storage device of claim 32 wherein the method steps
2 further comprise estimating a posterior probability $P(u|d,q)$ that the document d is
3 of interest to the user u , given a query q submitted by the user.

1 43. (*Original*) The program storage device of claim 42 wherein estimating the
2 posterior probability comprises estimating a probability $P(q|d,u)$ that the
3 query q is expressed by the user u with an information need in the document
4 d .

1 44. (*Original*) The program storage device of claim 32 wherein the method steps
2 further comprise applying the identified properties of the document d to a learning
3 machine having product parameters characterizing a product p to estimate a
4 probability $P(p|d)$ that the document d refers to the product p .

1 45. (*Original*) The program storage device of claim 44 wherein the method steps
2 further comprise updating the product parameters based on the identified
3 properties of the document d and the estimated probability $P(p|d)$.

1 46. (*Original*) The program storage device of claim 44 wherein the method steps
2 further comprise initializing the product parameters based on a set of
3 documents associated with the product p .

1 47. (*Original*) The program storage device of claim 32 wherein the method steps
2 further comprise clustering multiple users into clusters of similar users, wherein
3 the clustering comprises calculating distances between User Models, and selecting
4 similar users based on the calculated distances between User Models.

1 48. (*Original*) The program storage device of claim 32 wherein the method steps
2 further comprise calculating relative entropy values between User Models of
3 multiple users, and clustering together users based on the calculated relative
4 entropy values.

1 49. (*Original*) The program storage device of claim 32 wherein the parameters
2 defining the User Model comprise calculated distances between the User Model
3 and User Models of users similar to the user.

1 50. (*Original*) The program storage device of claim 32 wherein the method steps
2 further comprise selecting in a group of users an expert user in an area of
3 expertise, wherein selecting the expert user comprises finding an expert User
4 Model among User Models of the group of users, such that the expert User Model

indicates a strong interest of the expert user in a document associated with the area of expertise.

51. (*Original*) The program storage device of claim 32 wherein the method steps further comprise parsing the document d for hyperlinks, and separately estimating for each of the hyperlinks a probability that the hyperlink is of interest to the user u .

52. (*Original*) The program storage device of claim 32 wherein the method steps further comprise sending to a third party web server user interest information derived from the User Model, whereby the third party web server may customize its interaction with the user.

53. (*Original*) The program storage device of claim 32 wherein the monitored user interactions include a sequence of interaction times.

54. (*Original*) The program storage device of claim 32 wherein the method steps further comprise initializing the User Model using information selected from the group consisting of a set of documents provided by the user, a web browser history file associated with the user, a web browser bookmarks file associated with the user, ratings by the user of a set of documents, and previous product purchases made by the user.

1 55. (*Original*) The program storage device of claim 32 wherein the method steps
2 further comprise modifying the User Model based on User Model modification
3 requests provided by the user.

1 56. (*Original*) The program storage device of claim 32 wherein the method steps
2 further comprise providing to the user a score for a document identified by the
3 user, wherein the score is derived from the estimated probability.

1 57. (*Original*) The program storage device of claim 32 wherein the method steps
2 further comprise providing to the user a 3D map of a hyper linked document
3 collection, wherein the 3D map indicates a user interest in each document.

1 58. (*Original*) The program storage device of claim 32 wherein the method steps
2 further comprise temporarily using a User Model that is built from a set of
3 predetermined parameters of a profile selected by the user.

1 59. (*Original*) The program storage device of claim 32 wherein the method steps
2 further comprise initializing the User Model by selecting a set of predetermined
3 parameters of a prototype user selected by the user.

1 60. (*Original*) The program storage device of claim 59 wherein the method steps
2 further comprise updating the predetermined parameters of the prototype
3 user based on actions of users similar to the prototype user.

1 61. (*Original*) The program storage device of claim 32 wherein the method steps
2 further comprise identifying a set of users interested in the document *d*.

1 62. (*Original*) The program storage device of claim 61 wherein the method steps
2 further comprise calculating a range of interests in the document *d* for the
3 identified set of users.